DNAs and proteins or peptides specific to bacteria of the species *Neisseria meningitidis*, processes for obtaining them and their biological uses.

The invention relates to DNAs and to proteins and peptides which are specific to bacteria of the species *Neisseria meningitidis* (abbreviated below to Nm), to the process for obtaining them and to their biological uses, in particular for the prevention and detection of meningococcal infections and meningitis.

It is known that Nm is one of the main agents of cerebrospinal meningitis.

Studies conducted in the United States have shown that 5 to 10% of the population are asymptomatic carriers of the Nm strain(s). The transmission factors of Nm are poorly known. For a proportion of persons infected, Nm penetrates the bloodstream, where it can cause meningococcaemia and/or progress to the cerebrospinal stream, to cause meningitis. Without fast antibiotic treatment, the infection can develop like lightning and become fatal.

Compared with other pathogens, Nm has the characteristic of being able to cross the haemato-encephalic barrier to colonize the meninges. The study of the pathogenicity of Nm is therefore important not only in the context of meningitis, but also in the context of any disease which affects the brain.

The benefit of having available tools specific to this species of bacteria for the uses envisaged above is therefore understood.

Genetically, Nm is very close to bacteria of the species *Neisseria gonorrhoeae* (abbreviated to Ng below) and of the species *Neisseria lactamica* (abbreviated to Nl below). However, their pathogenicity is very different.

Nm colonizes the nasopharynx, and then crosses the pharyngeal epithelium to invade the submucous space, thus being responsible for septicaemia and meningitis.

Ng is especially responsible for infections located in the genitourinary tract. It colonizes the genital mucosa, and then crosses the epithelium, subsequently invading the subepithelium, where it multiplies and is responsible for a severe inflammatory reaction. Disseminated gonococcal infections are possible, but remain rare and are the result of only some strains.

As regards Nl, it is considered that this is a non-pathogenic strain, since it is not responsible for a localized or general invasion.

A first consideration thus led to taking into account the fact that Nm and Ng, while being bacteria very close to one another, have different pathogenic potencies.

Since the genome of these bacteria has a high homology, only limited parts of the genome of Nm and Ng must code for specific virulence factors responsible for their pathogenesis.

It is clear that Nm has, compared with Ng, DNA sequences which are specific to it and which must be involved in the expression of its specific pathogenic potency.

The species Nm is subdivided into serogroups based on the nature of the capsular polysaccharides.

At least 13 serogroups have been defined, among which serogroups A, B and C are responsible for about 90% of meningitis cases. Groups A and C are found in epidemic forms of the disease. Group B is the serogroup generally isolated the most in Europe and the United States.

The capsule, which is present in Nm and absent from Ng, has served as the basis for formulating meningococcal antimeningitis vaccines.

The polysaccharides of the Nm capsule have been used to formulate a vaccine which has proved to be effective in preventing in adults the meningitis caused by meningococci of serogroups A, C, W135 and Y.

However, the polysaccharide of Nm group C has proved to be weakly immunogenic in children of less than two years, while the polysaccharide of Nm group B is non-immunogenic in man and shares epitopes with adhesion glycoproteins present in human neuronal cells.

There is therefore no universal vaccine capable of preventing infections caused by all the serogroups of the meningococci and capable of responding to the intrinsic antigenic variability of bacterial pathogens in general and Nm in particular.

Because of the cross-reactivity of the Nm group B polysaccharide with human antigens, the multiplicity of the serogroups and the antigenic variability of Nm, the strategies proposed to date cannot lead to a vaccine which is effective in all situations.

Research is therefore concentrated on study of the characteristic elements responsible for the specificity of the meningococcal pathogenesis.

The majority of genes which have been studied in either of the two bacteria Nm or Ng have their homologue in the second bacterium.

In the same way, the majority of virulence factors identified to date in Nm have a counterpart in Ng, that is to say pilin, the PilC proteins, the opacity proteins and the receptors of lactoferrin and transferrin.

The specific attributes of meningococci characterized in the prior art are the capsule, the Frp proteins analogous to RTX toxins, Opc proteins of the external member, glutathione

peroxidase, the porin PorA and the rotamase gene.

Among these, only the capsule is invariably present in the virulent strains of Nm. However, several extracellular pathogens have a capsule without nevertheless crossing the haemato-encephalic barrier.

Attributes which have not yet been identified must therefore be responsible for the specificity of the meningococcal pathogenesis. These attributes are probably coded by DNA sequences present among the meningococci but absent from the gonococci.

The inventors have developed a new approach based on subtractive isolation of Nm-specific genes, which genes must be linked to the specific pathogenesis of Nm, and more particularly to crossing of the haemato-encephalic barrier.

The subtractive method developed in the prior art has resulted in the production of epidemological [sic] markers for some Nm isolates. These markers are of limited use: they do not cover all the serogroups of the Nm species.

In contrast to these studies, the work of the inventors has led, by confronting Nm with the entire Ng chromosome sheared in a random manner, to the development of a means for cloning all the DNAs present in Nm and absent from Ng, thus providing tools of high specificity with respect to Nm, and thus enabling the genetic variability of the species to be responded to for the first time.

The terms "present" and absent" used in the description and claims in relation to the DNAs of a strain or their expression products are interpreted on the basis of identical hybridization conditions (16 h at 65°C, with $NaPO_4$ 0.5 M, pH 7.2; EDTA-Na 0.001 M, 1%, 1% bovine serum albumin and 7% sodium dodecylsulphate) using the same probe and the same labelling intensity of the probe, the same amount of

chromosomal DNA and the same comparison element (chromosomal DNA of the homologous strain).

It is therefore considered that the DNA is present if the signal obtained with the probe is practically the same as that obtained with the reference strain.

Conversely, it is considered that the DNA is absent if this signal appears very weak.

A second consideration of the pathogenicities of Nm and Ng leads to taking into account their common capacity for colonization and penetration of the mucosa, and then invasion of the subepithelial space of the latter. It is highly probable that this process involves virulence factors common to the two pathogens. In this respect, it is known that a certain number of virulence factors have already been identified in Nm and in Ng, such as the pili proteins, PilC, the opacity proteins, the IgA proteases, the proteins for binding to transferrin and to lactoferrin, and the lipooligosaccharides.

The approach of the inventors is thus extended to investigation of the Nm regions which are specific to Nm and Ng but absent from the non-pathogenic species Nl, and in a general manner to investigation of the chromosomal regions of the DNAs and their expression products specific to a given species by the means developed in accordance with the invention.

The object of the invention is thus to provide DNAs of Nm specific to its pathogenic potency and means for obtaining them, in particular by formulating banks formed exclusively from these Nm-specific DNAs.

It also provides the products derived from these DNA sequences.

The invention also relates to the utilization of specific

and exhaustive characteristics of these banks to formulate tools which can be used, in particular, in diagnostics, treatment and prevention.

The DNAs of the invention are characterized in that they are in all or part genes, with their reading frame, present in *Neisseria meningitidis*, but absent both from *Neisseria gonorrhoeae* and from *Neisseria lactamica*, with the exception of genes involved in the biosynthesis of the polysaccharide capsule, frpA, frpC, opc, por A, rotamase, the sequence IS1106, IgA proteases, pilin, pilC, proteins which bind transferrin and opacity proteins.

As stated above, the terms "present" and "absent" are interpreted on the basis of the hybridization conditions used in the Southern blotting described in the examples and referred to above.

It can be seen that these DNAs include variants where these express a function intrinsic to the Nm species, more particularly a phenotype found solely in Nm or in common exclusively with Ng.

According to a main aspect, these DNAs are specific to the pathogenicity of *Neisseria meningitidis*, in spite of the genetic variability of this species.

According to an embodiment of the invention, the said DNAs are specific to Nm, in contrast to Ng.

More particularly, the Nm-specific DNAs are absent from *Neisseria lactamica* and from *Neisseria cinerea*.

Surprisingly, the majority of genetic differences between the strains of meningococci and those of gonococci appear grouped in distinct regions, which are said to correspond to the pathogenicity islets described previously for *E. coli* and *Y. pestis*.

In a preferred embodiment of the invention, these DNA are

thus also characterized in that they comprise one or more sequence(s) present on the chromosome of *Neisseria meningitidis* Z2491 between *tufA* and *pilT*, or region 1 of the chromosome, and/or the sequence(s) capable of hybridizing with the above sequence(s), with the proviso of being specific to *Neisseria meningitidis*.

"Specific" in the description and the claims means the nucleotide sequences which hybridize only with those of Nm under the hybridization conditions given in the examples and referred to above.

In this respect, it can be seen that, in a general manner, when "all or part" of a sequence is referred to in the description and claims, this expression must be interpreted with respect to the specificity defined above.

Furthermore, all or part of a peptide or a fragment of a peptide or an antibody means a product having the biological properties respectively of the natural peptide or the antibody formed against the peptide.

Genes of the *Neisseria meningitidis* capsule are grouped in region 1.

DNAs of this type have a sequence corresponding in all or part to SEQ ID No. 9, 13, 22 or 30, and/or to any sequence located at more or less 20 kb from these SEQ ID on the chromosome of an Nm strain, and/or have a sequence which is capable of hybridizing with at least a fragment of any one of these sequences.

In another preferred embodiment of the invention, these DNA are also characterized in that they are made up of one or more sequence(s) present on the chromosome of *Neisseria meningitidis* Z2491 between *pilQ* and λ740, or region 2 of the chromosome, and/or the sequences(s) capable of hybridizing with the above sequence(s), with the proviso of being specific

to *Neisseria meningitidis*.

DNAs according to this embodiment have a sequence corresponding in all or part to SEQ ID No. 1, 2, 4, 6, 7, 10, 15, 31 or 34, and/or to any sequence located at more or less 20 kb from these SEQ ID on the chromosome of an Nm strain, and/or have a sequence which is capable of hybridizing with at least a fragment of any one of these sequences.

The invention especially provides all or part of the DNA sequence SEQ ID No. 36 of 15,620 bp, and the sequences corresponding to the open reading frames SEQ ID No. 37, SEQ ID No. 38, SEQ ID No. 39, SEQ ID No. 40, SEQ ID No. 41, SEQ ID No. 42, SEQ ID No. 43, SEQ ID No. 44 and SEQ ID No. 45.

In yet another preferred embodiment of the invention, these DNAs are also characterized in that they are made up of one or more sequence(s) present on the chromosome of *Neisseria meningitidis* Z2491 between *argF* and *opaB*, or region 3 of the chromosome, and/or the sequence(s) capable of hybridizing with the above sequence(s), with the proviso of being specific to *Neisseria meningitidis*.

DNAs according to this embodiment are characterized in that they have a sequence corresponding in all or part to SEQ ID No. 8, 21, 23, 25, 26, 28, 29, 32 or 35, and/or to any sequence located at more or less 20 kb from these SEQ ID on the chromosome of an Nm strain, and/or have a sequence which is capable of hybridizing with at least a fragment of any one of these sequences.

Regions 1, 2 and 3 identified above have a high proportion of sequences specific to *Neisseria meningitidis* and also fall within the context of the invention.

Other DNAs representative of the specificity with respect to *Neisseria meningitidis* have one or more sequences which is/are present on the chromosome of *Neisseria meningitidis*

Z2491 but are not part of regions 1, 2 and 3 defined above.

Such DNAs comprise one or more sequence(s) corresponding in all or part to SEQ ID No. 3, 5, 11, 12, 14, 16, 18, 19, 20, 24, 27 or 33, and/or to any sequence located at more or less 20 kb from these SEQ ID on the chromosome of an Nm strain, and/or have a sequence capable of hybridizing with such sequences.

Taking into account the uses envisaged in particular, the invention more specifically relates to the above DNAs involved in the pathogenesis of the bacterial organism.

In particular, it provides the DNAs corresponding to at least one of the characterizations given above and coding for a protein exported beyond the cytoplasmic membrane, and/or of which all or part of their sequence corresponds to the conserved region of the said DNAs.

According to another embodiment of the invention, the DNAs are thus common with those of Ng, but are absent from those of Nl.

These are more specifically the DNAs which are present on region 4 (arg J to reg F) or on region 5 (lambda 375 marker to pen A) on the chromosome of Nm Z2491 and/or are capable of hybridizing with the said DNAs present, with the proviso of being specific to Nm and Ng, in contrast to Nl.

"Specific to Nm and Ng in contrast to Nl" means the DNAs which hybridize with the DNAs of Nm and Ng under the hybridization conditions of the examples (see example 4 in particular).

The DNAs of regions 4 and 5 and those capable of hybridizing with these DNAs, with the proviso of expressing the intrinsic functions of Nm, have the advantage of intervening in a significant manner in the virulence of Nm, being involved in the stage of initial colonization and

penetration and in the septicaemic dissemination.

According to other embodiments, the invention provides transfer and expression vectors, such as plasmids, cosmids or bacteriophages, comprising at least one DNA as defined above.

It also provides host cells transformed by at least one DNA as defined above.

Other host cells of the invention comprise genes or gene fragments specific to Nm, and are characterized in that their chromosome is deleted by at least one DNA according to the invention, in particular a DNA responsible for the pathogenicity. They are more specifically bacterial cells, in particular of Nm.

The invention also relates to the RNAs of which the sequence corresponds in all or part to the transcription of at least one DNA sequence or sequence fragment as defined above.

The invention also relates to the antisense nucleic acids of the DNAs as defined above, or of fragments of these DNAs.

These antisense nucleic acids carry, where appropriate, at least one substituent, such as a methyl group and/or a glycosyl group.

Other products which fall within the context of the invention include polypeptides.

These polypeptides are characterized in that they have an amino acid chain corresponding to all or part of a sequence coded by the nucleic acids defined above, or deduced from sequences of these nucleic acids.

They are advantageously polypeptides corresponding to all or part of the polypeptides exported beyond the cytoplasmic membrane, more specifically polypeptides corresponding to all or part of those coded by a conserved region.

As a variant, the polypeptides of the invention can be modified with respect to those corresponding to the nucleic

acid sequences such that they are particularly suitable for a given use, in particular use as a vaccine.

Modification is understood as meaning any alteration, deletion or chemical substitution where this does not affect the biochemical properties of the corresponding natural polypeptides, more specifically of functional proteins exported at the periplasm and the external membrane.

Other products according to the invention include antibodies directed against the above polypeptides.

The invention thus provides polyclonal antibodies, and also monoclonal antibodies, characterized in that they recognize at least one epitope of a polypeptide as described above.

It also relates to fragments of these antibodies, more particularly the fragments Fv, Fab and Fab'2.

The invention also relates to the anti-antibodies which are capable of recognizing the antibodies defined above, or their fragments, by a reaction of the antigen-antibody type.

According to the invention, the various products considered above are obtained by a synthesis and/or biological route in accordance with conventional techniques.

The nucleic acids can also be obtained from banks made up of Nm-specific DNAs such as are formulated by a subtractive technique, this technique comprising:

- mixing of two DNA populations,
- realization of at least one subtractive hybridization-amplification iteration, and
- collection of the desired DNA or DNAs, followed, where appropriate, by its/their purification with elimination of redundant sequences.

According to the invention, the two DNA populations originate respectively from a strain of *Neisseria*

meningitidis, the so-called reference strain for which the specific bank must be constructed, and a strain of *Neisseria*, the so-called subtraction strain, having a homology in primary DNA sequences of greater than about 70% with the *Neisseria meningitidis* strain, the DNA sequences of the subtraction and reference strains being obtained respectively by random shearing, and by cleavage by a restriction endonuclease capable of producing fragments less than about 1 kb in size.

The invention provides in particular a process for obtaining *Neisseria meningitidis*-specific DNA banks, comprising the stages of

- random shearing of the chromosomal DNA of a strain of *Neisseria gonorrhoeae*, the so-called subtraction strain, in particular by repeated passage through a syringe,
- cleavage of the chromosomal DNA of a strain of *Neisseria meningitidis*, the so-called reference strain, preferably by a restriction enzyme producing fragments less than about 1 kb in size,
- splicing of the DNA fragments of the reference strain, cleaved by the restriction enzyme, with suitable oligonucleotide primers,
- realization of a subtractive hybridization-amplification iteration, by:
  . mixing of the two DNA populations under suitable conditions for hybridization of homologous sequences, and then
  . amplification of auto-reannealed fragments and collection of these fragments,
  . digestion of these fragments by a restriction enzyme and re-splicing with oligonucleotide primers, followed by a
- purification of the spliced DNA and, where appropriate, a new iteration of the subtractive hybridization, comprising mixing of DNA fragments of *Neisseria gonorrhoeae* sheared as

indicated above with DNA fragments of *Neisseria meningitidis* produced by the preceding iteration, followed, if desired, by cloning of the DNAs of the bank.

The primers used are oligodeoxynucleotide primers which are suitable for the restriction endonuclease used and allow insertion into a cloning site, such as the EcoRI site of the plasmid pBluescript. Such primers will advantageously be chosen among the oligodeoxynucleotides referred to in the sequence listing under SEQ ID no. 36 to 45.

The banks thus obtained are formed from DNAs which are specific to meningococci and absent from gonococci.

The specificity of the DNAs was verified, as described in the examples, at each iteration by Southern blots, with genes common to the subtraction strain and to the reference strain, or with the total DNA of each of the strains digested by a restriction endonuclease, such as *Cla*I.

At each iteration, the exhaustivity of the DNA bank was also verified by Southern blotting with probes known to be specific to the reference strain, that is to say for *Neisseria meningitidis* the *frp, opc* and rotamase genes in particular.

The experiments carried out showed that the banks obtained by the process of the invention are deficient in genes having a significant homology with species of Neisseria other than *Neisseria meningitidis*, for example the *ppk* or *pil*C1 genes, generally in only 2 or 3 iterations.

If necessary, two routes, which are not exclusive of each other, can be taken.

It is possible to proceed with an $(n+1)^{th}$ iteration using the DNA of iteration n as the DNA population of the reference strain.

As a variant, a second bank independent of the first is constructed, with a restriction enzyme of different

specificity to that used in the first bank, for example *Mbo*I.

In all cases, it is preferable to keep each of the products produced by each of the iterations performed.

The invention also provides the use of the subtractive technique described above to obtain banks of the DNAs common to Nm and Ng, but specific with respect to Nl.

Three different banks are advantageously constructed, two of them by digestion of the chromosomal DNA of Nm by *MboI* and *Tsp5091*, and the third by digestion of the chromosomal DNA of Nm with *MspI*. Two subtraction series allow the DNAs having the required specificity to be collected, as described in the examples.

The invention also relates to the process for obtaining these banks and the banks themselves.

It can be seen that, generally, the process of the invention can be used to obtain banks of DNAs specific to a given cell species, or to a given variant of the same species, where another species or another variant which is close genomically and expresses different pathogenic potencies exists.

Using the process of the invention, DNA banks specific to given species of cryptococci, *Haemophilus*, pneumococci or also *Escherichia coli*, or more generally any bacterial agent belonging to the same species and having different pathovars will advantageously be constructed.

Furthermore, from these banks the invention provides the means to have available virulence factors specific to a species or a given variant.

Such banks are therefore tools which are of great interest for having available attributes which are responsible for the specificity of a pathogen, this use being more specifically illustrated according to the invention by the

obtaining of banks comprising the attributes responsible for the specificity of the meningococcal pathogenesis.

Study of the products of the invention, the nucleic acids, polypeptides and antibodies, has enabled an absolute specificity with respect ot *Neisseria meningitidis*, regardless of the strain and its variability, to be demonstrated.

These products are therefore particularly suitable for diagnosis or prevention of infections and meningitis caused by *Neisseria meningitidis*, whether in adults or children and regardless of the serogroups of the strain in question.

The method for diagnosis, according to the invention, of a meningococcal infection, and more particularly of meningococcal meningitis, by demonstration of the presence of *Neisseria meningitis* in an analytical sample is characterized by the stages of:

- bringing into contact a sample to be analysed, that is to say a biological sample or a cell culture, and a reagent formulated from at least one nucleic acid as defined above, if appropriate in the form of a nucleotide probe or a primer, or, as a variant, from at least one antibody or a fragment of an antibody as defined above, under conditions which allow, respectively, hybridization or a reaction of the antigen-antibody type, and
- detection of any reaction product formed.

If the reagent is formulated from a nucleic acid, this can be in the form of a nucleotide probe in which the nucleic acid or a fragment of the latter is labelled in order to enable it to be detected. Suitable markers include radioactive, fluorescent, enzymatic or luminescent markers.

As a variant, the nucleic acid is included in a host cell, which is used as the reagent.

In these various forms, the nucleic acid is used as such

or in the form of a composition with inert vehicles.

If the reagent is compiled from an antibody, or a fragment of an antibody, this can be labelled for detection purposes. Most generally, a fluorescent, enzymatic, radioactive or luminescent marker is used.

The antibody or the antibody fragment used, which is labelled if appropriate, can be used as such or in the form of a composition with inert vehicles.

The sample used in the stage of bringing the components into contact is a biological sample produced by a mammal, such as cephalorachidian fluid, urine, blood or saliva.

The detection stage is carried out under conditions which allow the reaction product to be demonstrated when it is formed. Conventional means use fluorescence, luminescence, colour or radioactive reactions, or also autoriadography [sic] techniques. It is also possible to quantify the product.

The invention also relates to the labelled products, the nucleic acids and antibodies, as new products.

The method defined above can be used for diagnosis of an immune reaction specific to a meningococcal infection.

The reagent used is thus a polypeptide according to the invention, as coded by the said nucleic acid sequences, corresponding to the natural product or a polypeptide which is modified but has the biological and immunological activity of the corresponding natural polypeptide.

It is advantageously a polypeptide exported beyond the cytoplasmic membrane of *Neisseria meningitidis*, more particularly the part of such a polypeptide corresponding to the conserved region of the DNA.

The invention also relates to kits for carrying out the methods defined above. These kits are characterized in that they comprise:

- at least one reagent as defined above, that is to say of the nucleic acid, antibody or polypeptide type,

- products, in particular markers or buffers, which enable the intended nucleotide hybridization reaction or immunological reaction to be carried out, as well as use instructions.

The specificity of the products of the invention and their location on the chromosome of *Neisseria meningitidis* Z2491, either grouped in a region and able to be interpreted as pathogenicity islets, or isolated on the chromosome, impart to them a very particular interest for realization of vaccine compositions with a universal purpose, that is to say whatever the strain and the variability which it expresses. These compositions can include in their spectrum other prophylaxes, and can be, for example, combined with childhood vaccines.

The invention thus provides vaccine compositions which include in their spectrum antimeningococcal prophylaxis, intended for prevention of any infection which may be caused by *Neisseria meningitidis*, these compositions being characterized in that they comprise, in combination with (a) physiologically acceptable vehicle(s), an effective amount of polypeptides or anti-antibodies or their fragments as defined above, these products optionally being conjugated, in order to reinforce their immogenicity [sic].

Immunogenic molecules which can be used comprise the poliovirus protein, the tetanus toxin, or also the protein produced by the hypervariable region of a pilin.

As a variant, the vaccine compositions according to the invention are characterized in that they comprise, in combination with (a) physiologically acceptable vehicle(s), an effective amount:

- of nucleic acids as defined above,

- of transformed host cells as defined above, or

- of Nm cells, the chromosome of which has been deleted by at least one DNA sequence according to the invention involved in the pathogenicity of the bacterium. The nucleotide material used is advantageously placed under the control of a promoter of its expression in vivo and synthesis of the corresponding protein. To reinforce the immunogenicity, it is also possible to combine this nucleic material with a DNA or an RNA which codes for a carrier molecule, such as the poliovirus protein, tetanus toxin or a protein produced by the hypervariable region of a pilin.

The vaccine compositions of the inventions can be administered parenterally, subcutaneously, intramuscularly or also in the form of a spray.

Other characteristics and advantages of the invention are given in the examples which follow for illustration thereof, but without limiting its scope.

In these examples, reference will be made to figures 1 to 11, which show, respectively,

- figures 1A, 1B, 1C, 1D, 1E, 1F and 1G: analysis of the subtractive bank *Tsp*5091,

- figure 2: the distribution of the Nm-specific sequences, in contrast to Ng, on the chromosome of the strain Z2491 (left-hand part) and of Nm-specific sequences, in contrast to Nl (right-hand part),

- figure 3A to 3C: the reactivity of the clones of the 3 regions of the chromosome according to the invention towards a panel of strains of the genus *Neisseria*,

- figure 4: the position in region 2 of the chromosome of Nm of oligonucleotides used as probes,

- figures 5, 6 and 7: the Southern blots of a panel of strains of the genus *Neisseria*, using parts of region 2 of Nm as

probes,

- figures 8A to 8C: the Southern blots with 3 subtractive banks over a panel of 12 strains of *Neisseria*, and

- figures 9, 10 and 11: the reactivity of clones of the 3 subtractive banks with respect to Nm, Nl and Ng.

In the examples which follow, the following materials and methods were used:

Bacterial strains - To obtain the subtractive banks, strain Z2491 of Nm (Achtman *et al.*, 1991, *J. Infect. Dis.* 164, 375-382), the strains MS11 (Swanson *et al.*, 1974, *Infect. Immun.* 10, 633-644) and the strains 8064 and 9764 of Nl were used, it being understood that any other strain of the species in question could be used.

In order to verify the specificity of these banks, 6 strains of Nm, 4 strains of Ng, one strain of Nl (*Neisseria lactamica*) and one strain of Nc (*Neisseria cinerea*) were used.

The six strains of Nm are: Nm Z2491 of serogroup A, Nm 8013 of serogroup C (XN collection), Nm 1121, no serogrouping possible (XN collection), Nm 1912 serogroup A (XN collection), Nm 7972 of serogroup A (XN collection) and Nm 8216 of serogroup B (XN collection).

The four strains of Ng are: Ng MS11 (Pasteur Institute, Paris), Ng 403 (Pasteur Institute, Paris), Ng 6934 (Pasteur Institute, Paris), Ng WI (isolated from a disseminated gonococcal infection), Ng 4Cl, Ng 6493 and Ng FA 1090.

The strains of Nl are Nl 8064 and Nl 9764 (XN collection), and that of Nc is Nc 32165 (XN collection).

Molecular genetics techniques

Unless indicated otherwise, the techniques and reagents used correspond to those recommended by Sambrook *et al* (Sambrook *et al* 1989, Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press). The

oligodeoxynucleotides used in this study are:

RBAm12, 3'AGTGGCTCCTAG 54 (SEQ ID No. 54)
RBam24, 5' AGCACTCTCCAGCCTCTCACCGAG 3'; (SEQ IN No. 55)
Jbam12, 3' GATCCGTTCATG 5'; (SEQ ID No. 60)
JBAM24, 5' ACCGACGTCGACTATCCATGAACG 3'; (SEQ ID No. 61)
REco12, AGTGGCTCTTAA; (SEQ ID No. 56)
~~REco24, 5'-AGCACTCTCCAGCCTCTCACCGAG-3';-(=-RBam-24)~~
JEco12, GTACTTGCTTAA; (SEQ ID No. 62)
JEco24, 5' ACCGACGTCGACTATCCATGAACG 3'; (= JBam24)
NEco12, AATTCTCCCTCG; (SEQ ID No. 64)
NEco24, AGGCAACTGTGCTATCCGAGGGAG; (SEQ ID No. 65).

Transfer to membranes (Southern blots)

The transfers to membranes were effected by capillary transfers to positively charged nylon membranes (Boehringer Mannheim). The hybridizations were carried out at 65°C in a solution comprising NaPi [sic] 0.5 M pH 7.2/EDTA 1 mM/SDS 7%/BSA 1%. The membranes were washed in a solution comprising NaPi [sic] 40 mM pH 7.2/EDTA 1 mM/SDS 1%. The final washing was carried out at 65°C for 5 min.

The probe *frp* obtained with oligonucleotides based on the *frp*A sequence corresponds to 2.4 kb of the 5' end of the gene of the strain Z2491. The *opc* and rotamase probes corresponding to whole genes are produced from the strain Z2491 using oligonucleotides constructed on the basis of published sequences. The probes *pilC1* and *ppk* (polyphosphate kinase) correspond to inserts of the plasmids pJL1 and pBluePPK6001 respectively.

Example 1: Construction of banks of DNAs present in Nm and absent from Ng.

a. "*MboI*" bank

Construction - The DNA of Nm Z2491 was cleaved by the endonuclease *Mbo*I and subjected to two iterations of a method called CDA (comprehensive difference analysis) below. This method comprises subtractive hybridization in the presence of excess sheared DNA of Ng MS11 and amplification by PCR of those meningococcal sequences which, since they are absent from or do not have significant homology with the DNA of Ng MS11, could reanneal

The chromosomal DNA of the strain Ng MS11 is sheared randomly by repeated passage through a hypodermic syringe until fragments of a size ranging from 3 to 10 kb are obtained. These DNA fragments are purified by extraction with phenol.

The chromosomal DNA of the strain Nm Z2491 is itself cleaved by the restriction endonuclease *Mbo*I. These DNA fragments (20 µg) are spliced with 10 nmol of annealed oligonucleotides RBam12 and RBam24. The excess primers are removed by electrophoresis over 2% agarose gel of low melting point. The part of the gel containing amplified fragments greater than 200 bp in size is excised and digested by β-agarase. These fragments are purified by extraction with phenol.

To carry out a subtractive hybridization (first iteration), 0.2 µg of the Nm DNA spliced with the RBam oligonucleotides is mixed with 40 µg Ng DNA in a total volume of 8 ml of a buffer EE 3X (a buffer EE 1X is composed of N-(2-hydroxyethyl)piperazine-N'-(3-propanesulphonic acid) 10 mM and EDTA 1 mM, and its pH is 8.0). This solution is covered with mineral oil and the DNA is denatured by heating at 100°C for 2 min. 2 µl NaCl 5 M are added and the mixture is left to hybridize at 55°C for 48 h. The reaction mixture is diluted to

1/10 in a preheated solution composed of NaCl and buffer EE, and in then immediately placed on ice.

10 µl of this dilution are added to 400 µl of PCR reaction mixture (Tris.HCl pH 9.0 10 mM; KCl 50 mM; $MgCl_2$ 1.5 mM; Triton X100 0.1%; 0.25 mM of each of the four triphosphate deoxynucleotides; Taq polymerase 50 units per ml). The mixture is incubated for 3 min at 70°C to complete the ends of the reannealed meningococcal DNA fragments.

After denaturing at 94°C for 5 min and addition of the oligonucleotide RBam24 in an amount of 0.1 nmol per 100 µl, the hybridizations are amplified by PCR (30 cycles of 1 min at 94°C, 1 min at 70°C and 3 min at 72°C, followed by 1 min at 94°C and 10 min at 72°C; Perkin-Elmer GeneAmp 9600).

The amplified meningococcal fragments are separated from the primers and high molecular weight gonococcal DNAs on gel. They are digested by *Mbo*I and the oligonucleotides JBam12 and JBam 24 are spliced to them again. These spliced DNAs are again purified over gel and extracted with phenol.

A second iteration of the subtractive hybridization is carried out on 40 µg of the randomly sheared Ng DNA and 25 ng of the DNA spliced with the JBam oligonucleotides obtained from the first iteration of the subtractive hybridization. During this second iteration, amplification of the auto-annealed Nm DNA is effected with the aid of the oligonucleotide JBam24.

Specificity - In order to confirm their Nm specificity, the amplified sequences after the second iteration of the CDA method are labelled and used as a probe for the DNA digested by *Cla*I produced from a panel of six strains of *Neisseria meningitidis*, four of *Neisseria gonorrhoeae*, one of *Neisseria lactamica* and one of *Neisseria cinerea*.

The Southern blots obtained show that the amplified

sequences resulting from the second iteration of the CDA method have a high reactivity with several bands corresponding to meningococci, and do not have a reactivity with the bands corresponding to the Ng, Nl and Nc strains.

The "MboI" bank thus appears to be Nm-specific.

Exhaustivity - In order to test the exhaustivity of the bank, all the products produced from the first and second iterations of the CDA method and also the initial chromosomal materials of Nm Z2481 [sic] and Ng MS11 are subjected to agarose gel electrophoresis, transferred to a membrane and brought into contact with probes comprising genes known to be meningococcus-specific, that is to say *frp*, *opc* and rotamase (Southern blotting).

As a result of these hybridizations, the Nm-specific gene *frp* is represented in the *Mbo*I bank by a fragment of 600 bp, but no activity is observed for the rotamase and *opc* genes. The *Mbo*I bank, although Nm-specific, therefore cannot be considered exhaustive.

Given their high specificity, the fragments produced by the second iteration of the CDA method for the *Mbo*I bank can nevertheless be cloned on the *Bam*HI site of the plasmid pBluescript.

A sequence corresponding to any of the Nm-specific genes can be included in the subtractive bank only if it is carried by a restriction fragment of appropriate size. This condition is a function of two factors. Firstly, the probability that the largest fragments are entirely Nm-specific is low. Secondly, even if such fragments existed, they would be under-represented in the bank because of the limitations of the PCR technique, the amplification effectiveness of which decreases with increasing size of the fragments. Fragments greater that about 600 bp in size are not included in the bank. Because of

the absence of *Mbo* fragments of suitable size from the chromosome of Nm Z2491, the rotamase and *opc* genes cannot be included in the bank. Any enzyme cannot by itself produce a small fragment corresponding to any Nm-specific gene. A second bank was therefore constructed using another restriction enzyme with a different specificity: *Tsp*509 [sic].

b. "Tsp5091" bank

Construction - The enzyme *Tsp*5091 has the advantage of producing fragments of smaller size (less than about 1 kb) than the enzyme *MboI*.

*Tsp*509I recognizes the sequence AATT and leaves, projecting at 5', a sequence of 4 bases compatible with *Eco*RI. The oligonucleotides used are Reco, Jeco and NEco.

The method followed conforms with that followed for construction of the "*Mbo*I" bank described above. However, higher quantities of meningococcal DNA were used for the first iteration of the subtractive hybridization in order to compensate for the higher number of fragments of low molecular weight produced by *Tsp*509I. For the first iteration, 400 ng Nm DNA fragments and, in the second, 25 ng Nm fragments are subjected to subtractive hybridization with 40 µg randomly sheared Ng DNA.

For the construction of this "*Tsp*509I" bank, as a control, a third iteration of the subtractive hybridization is carried out using 40 µg sheared Ng DNA and 0.2 ng Nm fragments resulting from a digestion by *Tsp*509I and a resplicing, with NEco adaptors, of the fragments obtained as a result of the second iteration.

Specificity - As described for the previous bank, the product resulting from the second iteration of the CDA method is labelled and used as the probe for a panel of strains of

Neisseria.

Figure 1A illustrates the Southern blot hybridization of products of the second iteration of the CDA method with the DNA digested by *ClaI* of: Nm in track a, Ng MS11 in track b, Nm 8013 in track c, Ng 403 in track d, Nm 1121 in track e, Ng 6934 in track f, Nm 1912 in track g, Ng WI (strain DGI) in track h, Nm 7972 in track i, Nl 8064 in track j, Nc 32165 in track k, Nm 8216 in track l.

In contrast to the high reactivity observed with all the Nm strains, a low or no reactivity is observed with the Ng, Nl and Nc strains.

The specificity of the bank was studied earlier by reacting membrane transfers (Southern blots) of the products produced by each of the three iterations of the CDA method with probes corresponding to *pil*C1 and *ppk*. These two genes are common to Nm and Ng.

Figure 1B shows an agarose gel after electrophoresis of the chromosomes of Nm Z2491 and Ng Ms11, digested by *Tsp*509 [sic], and products resulting from each of the iterations of the CDA method.

In track a 1 µg of the chromosome of Nm was deposited, in track b 1 µg of that of Ng, in track c 0.15 µg of the products resulting from the first CDA iteration, in track d 0.1 µg of those of the second iteration, in track e 0.05 µg of the third iteration, MW representing the molecular size markers.

Figures 1C and 1D show gels obtained as described in figure 1B after transfer to the membrane (Southern blots) and hybridization with *pil*C1 (figure 1C) and *ppk* (figure 1D).

At the end of the second iteration of the CDA method, the sequences corresponding to the *pil*C1 and *ppk* genes are completely excluded from the bank.

Exhaustivity - The exhaustivity of the bank was examined

by reacting the products resulting from the subtractive hybridization with the probes corresponding to three Nm-specific genes (*frp*, rotamase and *opc*).

These Nm-specific probes react with the amplification products resulting from the first and second iteration of the subtractive hybridization.

Figures 1E, 1F and 1G show gels obtained as described in figure 1B after transfer to the membrane (Southern blots) and hybridization with *frp*A (figure 1E), rotamase (figure 1F) and *opc* (figure 1G).

However, a third iteration of the subtractive hybridization leads to the loss of Nm-specific sequences, since the fragments which react with the rotamase and *opc* genes are absent from this third iteration.

In consideration of all these data, it emerges that the products resulting from the second iteration of the CDA method are Nm-specific and also constitute an exhaustive bank of Nm-specific sequences.

The products resulting from this second iteration are cloned at the *Eco*RI site of the plasmid pBluescript.

The bank produced by *Tsp*509I is more exhautive [sic] than the bank produced by *Mbo*I, as the theory considerations based on the enzymatic production of smaller restriction fragments would suggest.

In accordance with this aspect, it should be noted that the *Tsp*509I bank is less redundant than the *Mbo*I bank, that is to say it comprises less duplication of clones. 86% of the clones of the *Tsp*509I bank correspond to distinct sequences, while only 43% of the clones correspond to distinct sequences in the *Mbo*I bank (data not shown).

The bank produced by *Tsp*509I thus constitutes a source of Nm-specific clones.

Example 2: Analysis of the clones of the subtractive bank

Cloning and sequencing of the Nm-specific DNAs

The DNAs of the subtractive banks are clones at the *Bam*HI (*Mbo*I bank) or *Eco*RI (*Tsp*509I bank) site of the plasmid pBluescript, and then transformed in DH5α of *E. coli*. The inserts are amplified by PCR carried out on the transformed colonies using the primers M13-50 and M13-40, the latter primer being biotinylated on its 5' end.

Sequencing was carried out on each PCR product after separation of the biotinylated and non-biotinylated strands using the system of Dynabeads M-280 with streptavidin (Dynal, Oslo). The sequences are screened according to their homologies with previously published sequences using the computer programs Blastn and Blastx (NCBI, USA and Fasta).

The PCR products resulting from the transformed bacteria colonies after using the primers M13-40 and M13-50 as described above were labelled by incorporation with random priming of α-^{32}P-dCTP and were used as a probe for the membrane transfers of the chromosomal DNA digested by *Cla*I of strains Nm Z2491 and Ng MS11, as described above, in order to verify their specificity.

Mapping of clones on the chromosome of the strain Nm Z2491.

The results of studies carried out with 17 clones of the "*Mbo*I" bank (designated by the letter B) and 16 clones of the "*Tsp*5091" bank (designated by the letter E), each of these clones having a unique sequence and being without counterpart in Ng, are reported.

The positions of the DNA sequences corresponding to cloned Nm-specific products were determined with respect to

the published map of the chromosome of Nm Z2491 (Dempsey *et al.* 1995, J. Bacteriol. 177, 6390-6400) and with the aid of transfers to membranes (Southern blots) of agarose gel subjected to pulsed field electrophoresis (PFGE).

The Nm-specific clones are used as probes for a hybridization on membranes (Southern blots) of the DNA of Nm Z2491 digested with enzymes of rare cutting sites, that is to say *Pac*I, *Pme*I, *Sgf*I, *Bgl*II, *Spe*I *Nhe*I and *Sgf*I.

The gels (20 x 20 cm) were gels of 1% agarose in a buffer TBE 0.5X and were subjected to electrophoresis at 6 V/cm for 36 hours according to pulsation periods varying linearly between 5 and 35 seconds.

The hybridizations on the membrane (Southern blots) were carried out as described above.

The results obtained are shown on figure 2: the reactivity was located by comparison with the positions of the fragments of corresponding size on the published map. The positions of all the genetic markers mapped by Dempsey *et al* (mentioned above) are visualized with the aid of points on the linear chromosomal map. The Nm-specific genes disclosed previously are labelled with an asterisk. The two loci called "*frp*" correspond to the *frp*A and *frp*C genes. The "*pil*C" loci correspond to the *pil*C1 and *pil*C2 genes, which are pairs of homologous genes and are not distinguished on the map. The accuracy of the positions of the Nm-specific clones of the invention depends on the overlapping of reactive restriction fragments. On average, the position is +/- 20 kb.

This mapping reveals a non-random distribution of the Nm-specific sequences. The majority of the Nm-specific sequences belong to three distinct groups. One of these groups (region 1) corresponds to the position of genes relating to the capsule which have been described previously.

A distinction is made between:

- E109, E138, B230 and B323 as being region 1,

- B322, B220, B108, B132, B233, B328, E139, E145 as B101 as being region 2, and

- B306, E114, E115, E124, E146, E120, E107, E137 and 142 as being region 3.

63% of the sequences identified as specific to meningococci are located inside these three distinct regions.

This grouping contrasts with the distribution of previously disclosed Nm-specific genes (*frp*A, *frp*C, *por*A, *opc* and the region relating to the capsule).

This prior art would suggest in fact that the Nm-specific genes, with the exception of functional genes relating to the capsule, were dispersed along the chromosome.

Mapping of Nm-specific sequences on the chromosome leads to an unexpected result with regard to the prior art.

The majority of the genetic differences between the meningococcal and gonococcal strains tested are grouped in three distinct regions.

Meningococcal genes relating to the capsule are grouped in region 1.

The function of genes of the other regions is unknown, but homologies with published sequences (table 1) suggest similarities between certain genes of region 3 and bacteriophage transposase and regulatory proteins. No meningococcal virus has been characterized and it is tempting to think that these sequences are of phagic origin. Interestingly, the genome of H. *influenzae* also contains a sequence homologous to that of the Ner regulatory protein of phage Mu, but it is not known if it is a functional gene.

The clone B208 has a high homology (48% identical, 91% homology for 33 amino acids) with a clone of conserved regions

field III) in the class of proteins which bind to TonB-dependent ferric siderophors.

The proximity of this clone with the Nm-specific *porA* genes and the *frp* genes regulated by iron, and in particular the possibility that it is an Nm-specific receptor protein exposed on the external membrane in itself is a good candidate for further research.

The clone B339 corresponds to the Nm-specific insertion sequence IS*1106*.

The low homology between the clone B134 and the *Aeromonas* insertion sequence and also the presence of multiple copies of the clone B134 among the various strains of Nm suggest that it could be a new type of Nm-specific insertion sequence.

The possibility that the regions containing the Nm-specific clones could correspond to pathogenicity islets as described previously for *E. coli* and *Y. pestis* is of particular interest.

The clones isolated in this invention will allow better understanding of the relevance of Nm-specific regions in allowing cloning and sequencing of larger chromosomal fragments, and directly by their use for loci mutations.

Finally, detection of meningococcus-specific genes possibly involved in the pathogenicity of the organism allows targeting of suitable antigens which can be used in an antimeningococcal vaccine.

The effectiveness and the speed of the method according to the inventions enables it to be used in a large number of situations for which the genetic differences responsible for a phenotype peculiar to one of 2 close pathogens are investigated.

Study of the reactivity of the clones of regions 1, 2 and 3 towards a panel of strains of *Neisseria*.

The PCR products corresponding to inserts of each of the clones were collected and used as probes for hybridization on membranes (Southern blots) for a panel of strains of Nm, Ng, Nl and Nc.

Regions 1 and 2 produce a limited number of bands for each of the meningococci. This suggests that these regions are both Nm-specific and common to all the meningococci.

Figure 3 illustrates the reactivity of the clones of regions 1, 2 and 3 towards a panel of neisserial strains. The clones of regions 1 (figure 3A), 2 (figure 3B) and 3 (figure 3C) taken together were used as probes towards a panel of meningococci, gonococci and towards a strain of Nl and Nc.

The tracks are as follows: DNA of: Nm Z2491 in track a, of Ng MS11 in track b, of Nm 8013 in track c, of Ng 403 in track d, of Nm 1121 in track e, of Ng 6934 in track f, of Nm 1912 in track g, of Ng WI (strain DGI) in track h, of Nm 7972 in track i, of Nl 8064 in track j, of Nc 32165 in track k, and of Nm 8216 in track l.

Remarkably, region 3 has reactivity only with the meningococci of serogroup A. This region 3 is therefore specific to a sub-group of Nm.

A comparison was made with the known sequences in the databanks in order to evaluate the possible functions of the cloned regions.

Table 1 which follows gives the positions of specific clones on the chromosomal map and the homologies with known sequences.

TABLE 1 - Position of specific clones on the chromosomal map and homologies with known sequences

			Reactive fragments						
Name of clone*	Size of insert	Pac	Pmc	Bgl	Spe	Nhe	Sgf	Position on Z2491	Homologies of protein sequences
B305	259	18-20	15-17	22-23	18	11-13	2	λ736	
B333	235		15-17	22-23	18	11-13	2	λ736	
E109[1+]	211		6-7	11-15	10	11-13	2	*tufA* *ctrA*	protein LipB *N. meningitidis* (3×10^{-26})
E138[1+]	315	1	6-7	11-15	10	11-13	2	*tufA* *ctrA*	protein LipB *N. meningitidis* (4×10^{-75})
B230[1]	356	1-3	6-7	1	10	11-13	2	ctrA	protein KpsC *E. coli* (3×10^{-53})
B323[1]	363	1	6-7	1	10	11-13	2	ctrA	protein CtrB *N. meningitidis* (2×10^{64})
B322[2]	210		2	16-18	6	1	5	pilQ/λ 740	HlyB *S. marcescens* (4×10^{-15})

B220[2]	341		2	16-18	6	≥18	5	pilQ/λ 740	
B108[2]	275		2	19-21	6	>18	5	pilQ/λ 740	
B132[2]	411	2	2	19-21	6	≥18	5	pilQ/λ 740	
B233[2]	164	1-3	2	19-21	6	≥18	5	pilQ/λ 740	
B328[2]	256	1-3	2	22-23	6	≥18	5	pilQ/λ 740	
E139[2]	324	2	2	19-21	6	≥18	5	pilQ/λ 740	
E145[2]	343	2	2	19-21	6	≥18	5	pilQ/λ 740	
B101[2]	254	≥20	2	19-21	6	≥18	5	pilQ/λ 740	
E103q	334		2	11-15	3-5	10	3	λ644	
B326[§]	314		2	11-15	3-4	10	3	λ644	
B326 (low reactivity)			5	6	16	2	1	*argF*	
B342	167		2	19	3-4	6-7	3	*iga*	
E136	249		2	7	1	3	3	*lepA*	

B208	177		1	2	3-4	2	1	*porA*	FeIII pyochelin receptor *P. aeruginosa* (5.10^{-4})
= B306[3#]	219	11	5	11-12	5	2	4	parC	
E114[3]	227	11	5	11-12	5	2	4	parC	
E115[3#]	251		5	11-15	5	2	4	parC	
E124[3]	208		5	11-12	5	2	4	parC	
E146[3]	146		5	11-15	5		4	parC	
E120[3]	263		5	3-4	5	16	4	opaB	
E107[3]	248	11	14-17	3-4	5	16	4	opaB	
E137[3]	274		14-17	3-4	5	16	4	opaB	Transposase Bacteriophage D3112 (6×10^{-12})
E142[3]	230		14-17	3-4	5	16	4	opaB	Protein Ner-Likc *H. influenzae* (6×10^{-23}) Protein binding to the DNA Ner, phage mu (3×10^{-18})
E116	379	5-7	11-13	3-4	2	6-7	8	λ375	
B313	436	9	9	3-4	13-14	5	2	λ611	
B341	201	8-10	9	3-4	13-14	5	2	λ611	
E102	238		11-13	3-4	19	5	2	λ601	Hypothetical protein H11730 *H. influenzae*

B134	428			multi ple					(7 x 10^{-24}) transposase I*SAS2* *Aeromonas* *salmonicida* (5 x 10^{-5})
B339	259			multi ple					transposase IS *1106* *N. meningitidis* (6 x 10^{-45})

The level of homologies found, as given by the Blastx program, are indicated in parentheses

*) The clones labelled with the index "1", "2" or "3" belong to regions "1", "2" or "3" respectively of the chromosome of *N. meningitidis Z2491*.

+) E109 and E138 are contiguous clones §) B306 and E115 overlap #) B236 also has a low reactivity in the region of *arg* F

q) Clone E103 contains a *Tsp509* I site and can therefore contain two inserts; however, since it reacts only with a single fragment *Cla*I (Oks) of the chromosome of *N. meningitidis* Z2491 and occupies only one position on the map, this clone is included here.

Firstly, it can be seen that the clones of region 1 all correspond to genes involved in biosynthesis of the capsule. These genes have previously been studied among the Nm of serogroup B (Frosch *et al.* 1989, Proc. Natl. Acad. Sci. USA 86, 1669-1673 and Frosch and Muller 1993, Mol. Microbiol. 8 483-493).

With the exception of a low homology with the haemolysin activator of *Serratia marcescens*, the clones of region 2 have no significant homology with published sequences, either in the DNA or the proteins.

Two of the clones of region 3 have interesting homologies with proteins which bind to the DNA, in particular the bacteriophage regulatory proteins and transposase proteins.

Clone B208 has a high homology with one of the conserved regions in one class of receptors (TonB-dependent ferric siderophor).

Clones B134 and B339 hybridize with several regions of the chromosome (at least 5 and at least 8 respectively).

Data relating to the sequences show that clone B339 corresponds to the Nm-specific insertion sequence *IS1106*.

The translation of the clone B143 has a limited homology with the transposase of an *Aeromonas* insertion sequence (SAS2) (Gustafson *et al.* 1994, J. Mol. Biol. 237, 452-463). We were able to demonstrate by transfer on a membrane (Southern blots) that this clone is an Nm-specific entity present in multiple copies in the chromosomes of every meningococcus of the panel tested.

The other clones have no significant homology with the published neisserial sequences, and furthermore nor with any published sequence. These clones therefore constitute, with the majority of the other clones isolated, a bank of totally new Nm-specific loci.

Example 3: Study of region 2 of the Nm chromosome

. Determination and characterization of the sequence of region 2

PCR amplification is carried out with the chromosomal DNA of strain Z2491 of serogroup A, sub-group IV-1 using oligonucleotide primers formulated from each of the sequences of clones of region 2 in several different combinations. The PCR products which overlap are sequenced from the 2 strands using the chain termination technique and automatic sequencing (ABI 373 or 377).

To prolong the sequence beyond the limits of the clones available, partial SauIIIA fragments of 15 kb of the strain Z2491 are cloned in Lambda DASH-II (Stratagène).

The phages containing the inserts overlapping region 2 are identified by hybridization with clones of this region as probes. The DNA inserted is sequenced from the ends of the inserts, and these sequences are used to formulate new primers which will serve to amplify the chromosomal DNA directly, and not the phagic DNA.

An amplification of the chromosomal DNA is obtained using these new primers and those of the sequence previously available.

These PCR products are also sequenced from the 2 strands, which leads to a complete sequence of 15,620 bp (SEQ ID No. 36). The reading frames of this sequence which start with ATG or GTG and are characterized by a high codon usage index are analysed.

This analysis reveals 7 ORFs of this type which fill the major part of the sequence of 15,620 bp. The positions of these ORFs are the following:

ORF-1: 1330 to 2970 (SEQ ID No. 37); ORF-2: 3083 to 9025 (SEQ ID No. 38); ORF-3: 9044 to 9472 (SEQ ID No. 39); ORF-4: 10127 to 12118 (SEQ ID No. 40); ORF-5: 12118 to 12603 (SEQ ID No. 41); ORF-6: 12794 to 13063 (SEQ ID No. 43); ORF-7: 13297 to 14235 (SEQ ID No. 44); and ORF-8: 14241 to 15173 (SEQ ID No. 45).

ORF-4 starts with the codon GTG and overlaps a slightly smaller ORF (SEQ ID No. 41) in the same reading frame (9620-12118, frame 2), which starts with the codon ATG.

ORF-4 codes for a protein which has structural homologies with a family of polypeptides comprising pyocins (*Pseudomonas aeruginosa*), collcins and intimins (Escherichia coli), which are bactericidal toxins (pyocins, collcins) or surface proteins involved in adhesion of bacteria to eukaryotic proteins. ORF-7 encodes a protein, the sequence of which contains a potentially transmembrane region and which has structural homologies with periplasmic proteins or proteins inserted in the external membrane of bacteria. The structural homologies of ORF-4 and ORF-7 have been identified with the aid of the PropSearch program.

Investigation of sequences homologous to other ORFs in GenBank with the aid of the BLAST program revealed a homology between the N-terminal regions of ORF-2 and filamentous haemagglutinin B of *Bordetella pertussis* (43% similarity, 36% identical over 352 amino acids) and between ORF-1 and the protein fhaC of *Bordetella pertussis* (35% similarity, 27% identical over 401 amino acids). ORF-1 and ORF-2 are neighbouring genes in the strain Z249I and filamentous haemagglutinin B of *Bordetella pertussis* and fhaC are neighbouring genes in *Bordetella pertussis*, which reinforces the probability that these homologies reflect functional homologies.

. Confirmation of the specificity of region 2 with respect to Nm

Southern blots are carried out using the DNA probes obtained by PCR amplification of various parts of region 2 using oligonucleotide primers formulated from sequences of clones of region 2.

The approximate position of these oligonucleotides is shown on figure 4.

These are the oligonucleotides called R2001 (SEQ ID No. 46) and R2002 (SEQ ID No. 47) in one half of ORF-1, the oligonucleotides b332a (SEQ ID No. 48), e139a (SEQ ID No. 49), b132a (SEQ ID No. 50) and b233b (SEQ ID No. 51) in one half of ORF-1+the majority of ORF-2, and the oligonucleotides e145a (SEQ ID No. 52) and b101a (SEQ ID No. 53) in 1/3 of ORF-4+ORF-5 to 7.

The three Southerns are carried out under the following hybridization conditions:
16 h at 65°C,
$NaPO_4$ 0.5 M, pH 7.2
EDTA-Na 0.001 M
1% sodium dodecylsulphate.

For the washing, heating is carried out for 10 min at 65°C, and $NaPO_4$ 0.5 M, pH 7.2; EDTA-Na 0.001 M, 1% sodium dodecylsulphate are used.

Figures 5, 6 and 7 respectively show the Southern blots obtained with each of the abovementioned ORF parts.

The 14 tracks correspond respectively, in each of the Southerns, to
1: MS11 (Ng)
2: 403 (Ng)
3: FA1090 (Ng)

4: W1 (Ng)
5: 6493 (Ng)
6: marker (lambda hindIII)
7: Z2491 (Nm, gpA)
8: 7972 (Nm gpA)
9: 8013 (Nm, gpC)
10: 1121 (Nm, grouping not possible)
11: 1912 (Nm, gpB)
13: 32165 (Nc)
14: 8064 (Nl).

Given that a panel of strains of *Neisseria* is used in these experiments and that each well is charged with a similar amount of digested DNA, these 3 Southern blots clearly show that the sequences corresponding to region 2 are found in all the meningococci tested and that significant homologous sequences do not exist in the genome of the Ng of the strains tested.

Example 4: Identification of regions of the Nm genome absent from Nl and common with Ng

The technique described in example 1 is followed, but the chromosomal DNA of one strain of Nm (Z2491) and 2 strains of Nl (XN collections), equal parts of the DNAs of which are mixed, is used.

2 subtractions are performed using the R and J series of primers. Three different banks are thus obtained.

Two banks, called Bam and Eco, are obtained respectively by digestion of the chromosomal DNA of Nm Z2491 by *MboI* and *Tsp5091*; a third bank, called Cla, which results from digestion of the chromosomal DNA of Nm by *MspI*, is obtained

using the primer set RMsp10, RMsp24, JMsp10 and JMsp24. All the primers used are shown in the following table 2.

Table 2

Adapters for differential banks

Chromosomal DNA digested by Cloning in pBluescript by

*Mbo*I	→	*Bam*HI
*Tsp*509I	→	*Eco*RI
*Msp*I	→	*Cla*I

First subtraction cycle

RBam12 : 3' AGTGGCTCCTAG 5' (SEQ ID No. 54)
RBam24 : 5' AGCACTCTCCAGCCTCTCACCGAG 3' (SEQ ID No. 55)

REcol2 : AGTGGCTCTTAA (SEQ ID No. 56)
RBam24 : 5' AGCACTCTCCAGCCTCTCACCGAG 3' (SEQ ID No. 55)
(REco 24 = RBam 24)
RMsp10 : AGTGGCTGGC (SEQ ID No. 57)
RMsp24 : 5' AGCACTCTCCAGCCTCTCACCGAC 3' (SEQ ID No. 58)

Second subtraction cycle

Jbam12 : 3' GTACTTGCCTAG 5' (SEQ ID No. 59)
JBam24 : 5' ACCGACGTCGACTATCCATGAACG 3' (SEQ ID No. 60)

JEcol2 : GTACTTGCTTAA (SEQ ID No. 61)
JBam24 : 5' ACCGACGTCGACTATCCATGAACG 3' (SEQ ID No 60)

(JEco 24 = JBam 24)

JMsp10 : GTACTTGGGC (SEQ ID No. 62)
JMsp24 : 5' ACCGACGTCGACTATCCATGAACC 3' (SEQ ID No. 63)

After 2 subtractions, the entire product of each amplification is labelled and used as a probe.

The subtractive banks are checked by Southern blotting over a panel of 12 strains of *Neisseria* (chromosomal DNA cut by *ClaI*). The hybridization conditions are identical to those given in example 1.

These Southern blots are shown on figures 8A to 8C, which relate respectively to the *MboI/BamHI* bank, to the *MspI/ClaI* bank and to the *Tsp5091/EcoRI* bank.

The 12 tracks correspond respectively, to

1: Nm Z2491 (group A)
2: Nl 8064
3: Nm 8216 (group B)
4: Nl 9764
5: Nm 8013 (group C)
6: Ng MS11
7: Nm 1912 (group A)
8: Ng 4C1
9: Nm 1121 (grouping not possible)
10: Ng FA1090
11: Nc 32165
12: Nm 7972 (group A)

Examination of the Southern blots shows that the sequences contained in each bank are specific to Nm and are not found in Nl. Furthermore, the reactivity found with the strains of Ng suggests that some of these sequences are present in Ng.

Each of these banks was then cloned in pBluescript at the *BamHI* site for Bam, or the *EcoRI* sit for Eco, or the *ClaI* site for Cla. In order to confirm the specificity of the clones

with respect to the Nm genome, restriction of the individual clones and radiolabelling thereof were carried out. The clones showing reactivity for both Nm and Ng were kept for subsequent studies. These clones are shown on figures 9, 10 and 11, which give the profiles with respect to Nm, Nl and Ng of 5 clones of the Bam bank (figure 9), 16 clones of the Eco bank (figure 10) and 13 clones of the Cla bank (figure 11).

These clones were sequenced using universal and reverse primers. They are

- Bam clones

partial B11 of 140 bp (SEQ ID No. 66), partial B13 estimated at 425 bp (SEQ ID No. 67), B26 of 181 bp (SEQ ID No. 68), B33 of 307 bp (SEQ ID No. 69), B40 of 243 bp (SEQ ID No. 70),

- Cla clones

C16 of 280 bp (SEQ ID No. 72), partial C20 estimated at 365 bp (SEQ ID No. 73), partial C24 estimated at 645 bp (SEQ ID No. 74), partial C29 estimated at 245 bp (SEQ ID No. 75), C34 of 381 bp (SEQ ID No. 76), C40 of 269 bp (SEQ ID No. 77), C42 of 203 bp (SEQ ID No. 78), p C43 of 229 bp (SEQ ID No. 79), C45 of 206 bp (SEQ ID No. 80), C47 of 224 bp (SEQ ID No. 81), C62 of 212 bp (SEQ ID No. 82), and C130 (5'...) estimated at 900 bp (SEQ ID No. 83), and

- Eco clones

E2 of 308 bp (SEQ ID No. 84), partial E5 estimated at 170 bp (SEQ ID No. 85), partial E22 estimated at 300 bp (SEQ ID No. 86), E23 of 273 bp (SEQ ID No. 87), E24 of 271 bp (SEQ ID No. 88), E29 of 268 bp (SEQ ID No. 89), partial E33 estimated at 275 bp (SEQ ID No. 90), partial E34 estimated at 365 bp (SEQ ID No. 91), E45 of 260 bp (SEQ ID No. 92), E59 estimated at greater than 380 bp (SEQ ID No. 93), E78 of 308 bp (SEQ ID No. 94), E85 of 286 bp (SEQ ID No. 95), E87 of 238 bp (SEQ ID No. 96), partial E94 greater than 320 bp (SEQ ID No. 97), partial

E103 greater than 320 bp (SEQ ID No. 98) and E110 of 217 bp (SEQ ID No. 99).

Mapping of each clone was carried out on the chromosome of Nm Z2491 as described in example 1. The results obtained are given on the right-hand part of figure 2. It is found that these clones correspond to regions called 4 and 5. These regions are therefore made up of sequences present both in Nm and in Ng, but not found in Nl. It is therefore regarded that these are sequences which code for virulence factors responsible for the initial colonization and penetration of the mucosa. Region 4 is located between *argF* and *regF* on the chromosome of Nm 2491 [sic], and region 5 is located between the lambda 375 marker and *penA*. This region probably contains sequences which code for an Opa variant and a protein which binds transferrin.

A comparison with the known sequences in the databanks has half [sic] that in region 4 only the clone C130 has a homology, that is to say with *MspI* methylase. In region 5, no homology with known sequences was found with the clones C8, E2, B40, C45, E23 and E103. For the other clones, the homologies are the following:

B11 arginine decarboxylase SpeA; C29 arginine decarboxylase SpeA; C62 oxoglutarate/malate transporter; repetitive DNA element; E34 repetitive DNA element; E94 murine endopeptidase MepA ; C47 citrate synthase PrpC; E78 citrate synthase PrpC

Example 5: Demonstration of the presence of one or more strains of *Neisseria meningitidis* in a biological sample

A biological sample of the cephalorachidian fluid, urine, blood or saliva type is taken.

After filtration and extraction, the DNAs present in this

sample are subjected to gel electrophoresis and transferred to a membrane by Southern blotting.

A nucleotide probe constructed by labelling SEQ ID No. 5 with ^{32}P is incubated with this transfer membrane.

After autoradiography, the presence of reactive band(s) allows diagnosis of the presence of *Neisseria meningitidis* in the sample.

Example 6: Vaccine composition including in its spectrum antimeningococcal prophylaxis and intended for prevention of any form of infection by Neisseria meningitidis.

The peptide coded by a sequence including SEQ ID No. 10 is conjugated with a toxin.

This conjugated peptide is then added to a composition comprising the anti-*Haemophilus* and antipneumococcal vaccine, or any other childhood vaccine.

After having been sterilized, the resulting composition can be injected parenterally, subcutaneously or intramuscularly.

This same composition can also be sprayed on to mucosa with the aid of a spray.